V. METHODS USED TO RISK-ADJUST HOSPITAL MORTALITY DATA

Patients at different hospitals may vary in the severity of their pre-operative clinical condition. To make a fair comparison across hospitals, it is therefore necessary to adjust for differences in the risk-level of each hospital's patients. CCMRP "levels the playing field" by accounting for the pre-operative condition of each patient at the time he or she is admitted to the hospital. Hospitals that routinely handle "tougher" cases get a larger risk-adjustment factor, while hospitals that handle "easier" cases get a smaller factor. Note that CCMRP intends to include as risk-adjustment variables only those data elements that describe the patient's condition as closely as possible to the time of hospital admission. The goal is to produce a statistical model that can be used to risk-adjust hospital outcomes by removing patient factors existing prior to the hospitalization that can affect survivorship.

The text below summarizes the methods used to risk-adjust hospital mortality data. Readers interested in a more thorough explanation of the data, risk-adjustment methods, and results should refer to Appendix F.

Data

The risk analysis is based on 30,814 isolated CABG cases for 82 California hospitals that submitted data to CCMRP for 1997 and 1998. Data for these 82 hospitals represent more than 70% of the isolated CABG cases performed in California.[8] CCMRP collected a set of 41 data elements for each patient who underwent an isolated CABG procedure at the participating hospitals. The data elements (listed in Table 1) focus on demographic characteristics and the pre-operative condition—also known as risk factors of the patient. The outcome measure utilized was **in-hospital mortality** (i.e., the deaths that occurred in the same hospital admission).

CCMRP evaluated the data submitted from each hospital for completeness and potential data errors. When problems were identified, CCMRP contacted hospital staff to encourage investigation of potential data errors, and, when necessary, to request replacement of incomplete or erroneous data. When data were missing from the hospital submission, CCMRP replaced the blank field with the lowest risk-category for the variable that was missing. For example, if the hospital left the field *diabetes* (yes or no) unmarked, CCMRP presumed the condition was not present for that patient and assigned a "no" to that field. Likewise, if the field for *NYHA congestive heart failure class* was unmarked, we assigned the lowest risk category to this record—in this case, NYHA Class I. The CCMRP policy decision to assign the lowest risk category to any missing data element was based on three factors: 1) many hospitals may leave data fields blank by design (e.g., blank means a comorbid condition was not present or was a STS coding convention such as for creatinine <2.0); 2) consistency with the other major cardiac reporting programs, which recode missing data with the lowest or normal value; and 3) declining to give hospitals any additional credit in the risk model when coding is incomplete, thereby creating an incentive for more complete coding.

[8] Three of the 82 Hospitals that submitted data for the 1997-1998 period withdrew from the program after the analysis was completed but prior to preparation of the report, leaving 79 hospitals that agreed to publicity report their results. However, data from all 82 hospitals was used to develop the risk-adjustment model.

After preliminary data cleaning and analyses were completed, CCMRP developed and implemented an audit process designed to review the quality of the data submitted for 1998. The intent of the audit was to determine whether the rating received by the hospital was in any way a function of that hospital's coding practices. That is, did hospitals classified as better performers systematically overstate the severity of their cases, or did hospitals classified as worse performers systematically understate the severity of their patient case-mix? Twenty-six hospitals were audited out of the 79 that are publicly reporting for the first round of data collection, or 33% of the hospitals reporting. CCMRP concluded from the audit analysis that there was no relationship between a hospital's average patient risk-level and the rating received by the hospital.

Risk Model

CCMRP used a multivariate logistic regression model to determine the relationship between each of the demographic and pre-operative risk variables and the likelihood of in-hospital mortality. Multivariate logistic regression models relate the probability of death to the explanatory factor, (e.g., patient age, the amount of creatinine in the blood, or the anginal status of the patient) while controlling for all other explanatory factors in the model. For example, the odds ratio of 1.05 for age derived in CCMRP model means that a patient one year older than another will have an odds of dying 1.05 times higher—when all other factors are held constant. Table 4 presents the final model based on the 1997-1998 CCMRP data set.

Table 4: CCMRP 1997–1998 Logistic Regression Model

Explanatory Factor	Coefficient	Std. Error	t–value	Odds Ratio	Missing Variable Assignment
(Intercept)	–7.206	0.411	–17.512		
Age (in years)	0.044	0.004	**10.812**	1.05	Case Excluded
Sex					
Female	Reference				
Male	–0.401	0.080	**–5.005**	0.67	Male
Race					
White	Reference				White
Non–white	0.203	0.088	**2.294**	1.23	
Creatinine (mg/dl)	0.214	0.039	**5.433**	1.24	1.0; Truncated at 10
Hypertension	0.075	0.087	0.866	1.08	No
Dialysis	–0.029	0.275	–0.105	0.97	No
Diabetes	0.142	0.080	1.776	1.15	No
Peripheral Vascular Disease	0.435	0.091	**4.800**	1.54	No
Cerebrovascular Disease	0.244	0.101	**2.410**	1.28	No
Ventricular Arrhythmia	0.337	0.123	**2.737**	1.40	No
COPD	0.275	0.094	**2.914**	1.32	No
Operative Incidence					
First	Reference				First Operation
Second	0.674	0.118	**5.733**	1.96	
Third	1.354	0.276	**4.901**	3.87	
Fourth or Higher	1.823	0.660	**2.763**	6.19	
Myocardial Infarction					
None	Reference				None
Yes, but When Unknown	0.156	0.196	0.797	1.17	
21+ Days ago	0.028	0.105	0.263	1.03	
7–20 Days ago	–0.227	0.198	–1.145	0.80	
1–6 Days ago	0.237	0.107	**2.211**	1.27	
Within 1 day	0.876	0.150	**5.831**	2.40	
PTCA on This Admission	0.220	0.156	1.411	1.25	No
Angina					
None	Reference				
Stable	–0.369	0.137	**–2.691**	0.69	Angina Stable
Unstable	–0.256	0.129	**–1.977**	0.77	
NYHA CHF Class					
I	Reference				
II	0.506	0.122	**4.141**	1.66	NYHA Class I
III	0.549	0.109	**5.037**	1.73	
IV	0.769	0.102	**7.530**	2.16	
CCS Angina Class					
I	Reference				
II	0.178	0.192	0.927	1.19	
III	0.070	0.173	0.404	1.07	CCS Class III
IV	0.211	0.175	1.203	1.23	

Table 4: CCMRP 1997–1998 Logistic Regression Model (cont.)

Explanatory Factor	Coefficient	Std. Error	t–value	Odds Ratio	Missing Variable Assignment
Acuity					
Elective	Reference				Elective
Urgent	0.221	0.090	**2.449**	1.25	
Emergent	0.743	0.136	**5.482**	2.10	
Salvage	2.806	0.218	**12.860**	16.55	
Ejection Fraction (%)	–0.012	0.003	**–4.393**	0.99	55; Truncated at 15.0
Left Main Stenosis					
0–50%	Reference				0–50%
51–70%	–0.015	0.126	–0.117	0.99	
71–90%	0.233	0.130	1.786	1.26	
91+%	0.525	0.153	**3.426**	1.69	
Type of Coronary Disease					
Single Vessel	Reference				Single Vessel Disease
Double vessel	–0.176	0.181	–0.974	0.84	
Triple or More	0.069	0.160	0.433	1.07	
LM Only disease	0.447	0.359	1.244	1.56	
Mitral Regurgitation					
None	Reference				None
Trivial	0.506	0.158	**3.203**	1.66	
Mild	0.247	0.151	1.638	1.28	
Moderate	0.612	0.192	**3.187**	1.84	
Severe	0.898	0.345	**2.598**	2.45	

Age, ejection fraction, and creatinine were entered as continuous variables; the other variables were entered as ordered factors. For the variables entered as ordered factors, the coefficients should be compared to the reference category (for example, we show coefficients for NYHA Classes II, III, and IV; those coefficients are compared to the reference category of NYHA Class I). Bolded t–values indicate the coefficient for that variable is statistically significant at the 0.05 level.

GUIDE TO INTERPRETING THE RISK MODEL

Coefficient: The coefficient of the explanatory factor indicates the effect of a patient having the characteristic on the likelihood of in-hospital death following bypass surgery. If the value is positive, it means that the characteristic is associated with an increased risk of death compared to not having the characteristic—while controlling for the effect of all of the other factors. If the coefficient is negative, having that characteristic is associated with a lower risk of death compared to not having it. The larger the value (whether positive or negative), the greater the effect or weight this characteristic has on the risk of dying. For example, note that the coefficient for *peripheral vascular disease (PVD)* is 0.435 and significant. This value is positive, so it indicates that CABG patients with *peripheral vascular disease* are at an increased risk of dying in the hospital compared to patients that do not have the disease. On the other hand, the coefficient for the variable *male* has a value of –0.401. Since the value is negative, it means that males have a lower probability of dying in the hospital than females—after taking into account all other factors.

Standard Error: The standard error is the standard deviation of the sampling distribution of an estimate, and is a measurement of the statistical reliability of that estimate. The coefficient divided by the standard error produces the t-statistic.

t–Value: The t-value is a measure of the statistical significance of the coefficient. When the t-value is large (whether positive or negative), it means that we are relatively confident that the effect of the factor is real. If the t–value is small, we are less confident that the effect was not observed by chance alone. A common rule of thumb for interpreting this column is that if the absolute t-value is larger than 2.0, we have some confidence that the effect of the factor is real. For example, the t-value for the male explanatory factor is -5.005. Since it's absolute value is greater than 2.0, we have some confidence that the sex of the patient is a statistically significant factor in explaining in–hospital mortality for CABG patients. **Not all of the explanatory factors in our model have t-values that are larger than 2.0.** For example, the t–values for CCS angina class and type of coronary artery disease (single vessel disease, double, triple, or left main only disease) are all quite small. This indicates that, for our data, neither coronary disease type nor CCS class are reliable predictors of in-hospital mortality. Note that a small t–value does not mean that factor has no effect on in-hospital mortality—it means that the effect, if any, is not reliably estimated.

Odds Ratio: Another way of assessing the impact of each factor on in–hospital mortality is to utilize the odds ratio. Mathematically, the odds ratio is simply the antilogarithm of the coefficient value, but it is often easier to interpret. The larger the odds ratio, the greater the impact that characteristic has on the risk of dying. An odds ratio close to 1.0 means that the effect of the factor is close to neutral. For example, the odds ratio for *peripheral vascular disease* (PVD) is

1.54. This means that if the patient has *peripheral vascular disease* the odds of dying in–hospital are about 1.54 times higher than if the patient did not have PVD. Being *male* has an odds ratio of 0.67, which means that the odds that a man will die in–hospital after CABG surgery is about 0.67 times as high (i.e., about two thirds as much) as for a woman.

Missing Data Assignment: When data were missing from the hospital submission, CCMRP replaced the blank fields with the lowest risk category for the variable that was missing. For example, if the hospital left the field for *NYHA congestive heart failure class* unmarked, we assigned the lowest risk category to this record—in this case, NYHA Class I. This column indicates the specific category used to replace missing data for each variable.

Key Technical Findings Regarding the Risk Model

- Although several of the variables do not appear to be "statistically significant" (as determined by the t–value), almost all coefficients appear with the expected sign from a clinical standpoint.
- Age, acuity (i.e., how urgent the operation was), ejection fraction, and operative incidence are very important risk–model variables.
- Even after controlling for all other variables, sex appears to have a statistically significant effect, with males having about one–third lower mortality. The literature suggests that sex may serve as a proxy for body size; unfortunately, although the CCMRP attempted to collect height and weight to construct an index of body mass, the analysis was hampered by missing values and the apparent confusion of metric (kilogram and centimeter) and English (pound and inch) units in the data submission.
- After accounting for creatinine levels, dialysis appears to have no additional explanatory power. That is, given that a dialysis patient has higher creatinine levels than the average patient, once one knows that level, the fact that the patient is on dialysis appears to add no additional information.
- Patients with no angina have higher risk of in–hospital death than patients reported as having either "stable" or "unstable angina." Patients with no angina are unusual in that the majority of patients undergoing isolated CABG surgery have either "stable" or "unstable angina." Table F–1 (**Technical Appendix**) shows that only about 10% of the patients are classified as having "angina, none."
- The New York Heart Association (NYHA) Class, used to measure the severity of congestive heart failure, appears to make a "natural" split between NYHA Class I and NYHA Classes II, III, and IV.
- Canadian Cardiovascular Society (CCS) Class, used to measure the severity of angina, does not appear to have much explanatory power. Since the majority of CABG patients suffer from Class III or Class IV anginal pain, there is probably insufficient variability in these data to distinguish mortality differentials.

- The coefficients on the Myocardial Infarction (MI) variable seem to indicate that an MI more than one week before the CABG procedure has an effect on risk indistinguishable from no MI at all, even after controlling for the acuity of the operation.
- Moderate amounts of stenosis of the Left Main coronary artery (up to about 70% stenosis) do not appear to have a significant elevating effect on the risk of in–hospital mortality. Stenosis beyond the 70% level appears to have a much larger effect. Note that the usual analysis might conclude that a 75% stenosis is statistically indistinguishable from no stenosis because the t–statistic is less than 2.0 (it is 1.78).[9]
- Among the collected comorbidities, peripheral vascular disease appears to have the largest effect.
- The number of vessels affected with coronary disease appears to have an effect in the hypothesized direction. The risk of death increases (ie., with greater a number of vessels affected), but the effect is not statistically distinguishable from no effect.
- While "moderate" and "severe" mitral regurgitation appear to have effects as would be expected from a clinical standpoint, "mild" regurgitation is anomalous in appearing to have a lesser effect than "trivial." This may result from coding confusion between these two categories and CCMRP intends to focus on this distinction in future data collection training sessions.

[9] For the year 2000, the STS Adult Cardiac Database will be collecting data only on whether stenosis of the left main coronary artery exceeds 50% and will no longer collect data on the degree to which stenosis is beyond 50%.